

Speech recognition system as means of controlling access

AB

(DE19522940)

The voice controlled access system has a memory 2, a processor 1, a recorder 4, a loudspeaker 8, several batteries 3,6, two input keys T1,T2 and a main switch 5 controlling the activation of the vehicle electronic system 7. A spoken command is recorded, and the analogue signal is converted into a digital format for entering into the memory. The processor compares the command with the digitised reference speech patterns.

IN WAGNER THOMAS BAUER NORBERT

PA FRAUNHOFER

PA0 Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V., 80636 München, DE

Published As

	Publ. number	Pub. date	Appl. number	Appl. date	Publ. Stage
	DE19522940	19970102	1995DE-1022940	19950623	A1 - Doc. laid open (First publication)

PR

1995DE-1022940 19950623

IC

B60R-025/00
B60R-025/04
G07C-009/00
G10L-007/08
G10L-017/00

ICAA

B60R-025/00 [2006-01 A - I R M EP];
G07C-009/00 [2006-01 A - I R M EP];
G10L-017/00 [2006-01 A - I R M EP]

ICCA

B60R-025/00 [2006 C - I R M EP];
G07C-009/00 [2006 C - I R M EP];
G10L-017/00 [2006 C - I R M EP]

EC

B60R-025/00
G07C-009/00C2D
G10L-017/00U

CT

(DE19522940)

Search Report [Examiner]

DE4027491(A1) [DE4027491]

DE9419006(U1) [DE9419006U]

US5394135(A) [US5394135]

WO9215090(A1) [WO9215090]



19 BUNDESREPUBLIK
DEUTSCHLAND



DEUTSCHES
PATENTAMT

12 **Offenlegungsschrift**
10 **DE 195 22 940 A 1**

51 Int. Cl.⁸:
G 10 L 7/08
G 07 C 9/00
B 60 R 25/00
B 60 R 25/04

21 Aktenzeichen: 195 22 940.1
22 Anmeldetag: 23. 8. 95
43 Offenlegungstag: 2. 1. 97

DE 195 22 940 A 1

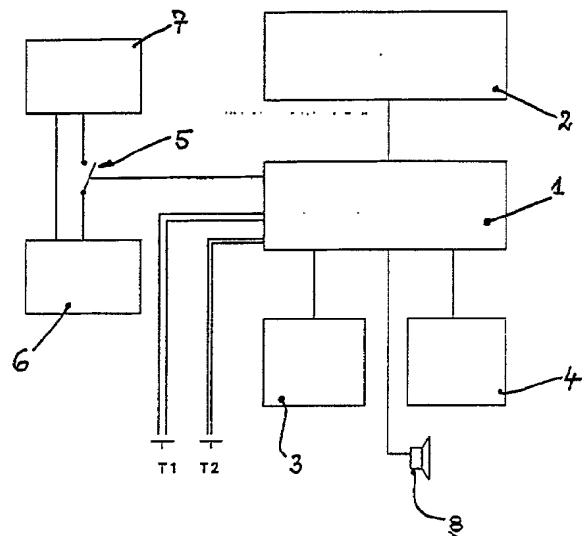
71 Anmelder:
Fraunhofer-Gesellschaft zur Förderung der
angewandten Forschung e.V., 80636 München, DE
74 Vertreter:
Leonhard und Kollegen, 80331 München

72 Erfinder:
Wagner, Thomas, 91052 Erlangen, DE; Bauer,
Norbert, 91058 Erlangen, DE
56 Entgegenhaltungen:
DE 40 27 491 A1
DE 94 19 006 U1
US 53 94 135
WO 92 15 090 A1

Prüfungsantrag gem. § 44 PatG ist gestellt

54 Verfahren zum Identifizieren einer menschlichen Stimme

57 Die Erfindung stellt ein zuverlässig arbeitendes Verfahren und eine solche Vorrichtung zum Identifizieren einer menschlichen Stimme bereit, insbesondere zum Sichern von Gegenständen gegen Diebstahl, umfassend
(a) das Aufnehmen der Stimme mit Hilfe einer Aufnahmeeinheit,
(b) das Umwandeln der aufgenommenen akustischen Signale in elektrische Signale,
(c) das Digitalisieren und Abspeichern der Signale in einer Auswerteeinheit mit Speicher, wobei ggf. die obigen Schritte (a) bis (c) mit weiteren Stimmen wiederholt werden und
(d) das Vergleichen der digitalisierten Signale der Stimme mit solchen, die sich bereits im Speicher der Auswerteeinheit befinden, und ggf. das Identifizieren der Stimme als im wesentlichen identisch mit einer der Stimmen, die im Speicher der Auswerteeinheit gespeichert sind.



DE 195 22 940 A 1

Die vorliegende Erfindung betrifft das Gebiet der akustischen Personenidentifikation. Das erfindungsgemäße Verfahren eignet sich beispielsweise zum Sichern von Gegenständen gegen Diebstahl, wobei die Identifikation einer Stimme als Schlüssel fungiert. Diese Identifikation muß zuverlässig arbeiten, sollte aber gleichwohl im Rahmen eines vertretbaren Hard- und Softwareaufwandes liegen.

Erfindungsgemäß wird diese Aufgabe dadurch gelöst, daß ein Verfahren zum Identifizieren einer menschlichen Stimme bereitgestellt wird, welches das Aufnehmen der Stimme mit Hilfe einer Aufnahmeeinheit, das Umwandeln der aufgenommenen akustischen Signale in elektrische Signale sowie das Digitalisieren und Abspeichern der Signale in einer Auswerteeinheit mit Speicher umfaßt. Erfindungsgemäß ist es möglich, daß mehrere Stimmen und/oder mehrere Laut folgen ein und derselben Stimme mittels der genannten Schritte im Speicher abgelegt, d. h. "eingelernt" werden. Zum Identifizieren einer bestimmten Stimme werden dann die wie oben erhaltenen digitalisierten Signale dieser Stimme mit solchen anderer Stimmen, die sich bereits im Speicher der Auswerteeinheit befinden, verglichen. Dabei kann die Stimme vom System als im wesentlichen identisch mit einer der Stimmen, die bereits gespeichert sind, identifiziert oder als unbekannt erkannt werden (Anspruch 1, Anspruch 10).

Das erfindungsgemäße Verfahren läßt sich besonders vorteilhaft beim Sichern von Gegenständen einsetzen, wobei die Identifikation der Stimme als "Schlüssel" fungiert (Anspruch 11). Dieser "Schlüssel" kann beispielsweise das Unterbrechen oder Schließen eines Stromkreises sein. Auch andere Arten von Schließfunktionen sind denkbar (Computer-Zugangsberechtigung, Personen-Identifikation beim Online-Banking oder Einsatz in Geldautomaten). Um einen solchen Freigabe- oder Schließmechanismus zu konstruieren, wird eine Stimme oder werden mehrere Stimmen, bevorzugt in Form einer charakteristischen Laut- bzw. Tonabfolge, in digitalisierter Form im Speicher abgelegt. Die digitalisierten Charakteristika dieser Stimme oder Stimmen dienen als "Schloß". Nur solche Personen, deren Stimmen abgespeichert sind, verfügen über den zu diesem Schloß gehörigen "Schlüssel". Zum "Aufschließen" muß diese Person die genannten Laut- oder Tonfolgen, beispielsweise ein bestimmtes Codewort, von sich geben. In einer bevorzugten Ausführungsform der Erfindung läßt sich diese "Schließfunktion" aktivieren und deaktivieren, beispielsweise indem zwei Schalter den Ablauf der einzelnen Verfahrensschritte so steuern, daß der eine Schalter das Aufnehmen, Umwandeln und Speichern der Signale ermöglicht, während der andere Schalter das Vergleichen und ggf. Identifizieren ermöglicht (Anspruch 2).

In einer bevorzugten Ausführungsform läßt sich das Verfahren zum Sichern von Kraftfahrzeugen gegen Diebstähle einsetzen (Anspruch 4, Anspruch 5). Bekanntlich nimmt die Zahl von Kfz-Diebstählen, insbesondere bei teuren Modelle, drastisch zu, so daß erhöhte und zusätzliche Schutzmaßnahmen erforderlich sind. Gleichzeitig setzen Versicherungen für eine Kostenerstattung im Schadensfall wirksame Sicherungsmaßnahmen gegen unbefugte Benutzung voraus. Besonders vorteilhaft ist das vorliegende Verfahren, weil es mit einem unverlierbaren Schlüsselsystem arbeitet. Sicherungssysteme im Stand der Technik basieren im wesentlichen auf mechanischen oder elektronischen Schlüsseln,

ohne die zwar eine unbefugte Benutzung oft wirkungsvoll verhindert wird, bei deren Verlust oder Diebstahl das Kfz jedoch sofort entwendet werden kann.

In einer bevorzugten Ausführungsform wird bei der Verwendung des erfindungsgemäßen Verfahrens zum Sichern von Kraftfahrzeugen deren Elektrik unterbrochen und erst dann wieder freigegeben, wenn sich ein Benutzer durch Sprechen eines Codewortes, beispielsweise seinen Namen, oder aber auch durch ein anderes

eingespeichertes akustisches Signal zu erkennen gibt. Ausschlaggebend für die Identifikation ist dabei die Kombination vieler Merkmale der Stimme der jeweiligen Person, beispielsweise Stimmfarbe, Tonhöhe, usw. die sich von anderen nicht imitieren lassen; ähnlich der Eigenschaft von Fingerabdrücken.

Die Verwendung des erfindungsgemäßen Verfahrens ist auch deshalb besonders vorteilhaft, da Speicher- und Logikbausteine ständig billiger werden und die meist bereits vorhandene Elektronik im Kfz, insbesondere bei teureren Modellen, beispielsweise das Vorhandensein eines Bordcomputers mit Speicher und Recheneinheit, die Installation des entsprechenden Systems kostengünstig ermöglicht, bei schon vorhandenen Hardware-Konzept und Ergänzung eines vertretbaren Aufwandes an Software.

Um den Gegenstand zu sichern, ist die Auswerte- und Steuereinheit, mit der das Aufnehmen und Vergleichen der Stimmen ermöglicht wird, so mit dem zu sichernden Gegenstand verbunden, daß die Benutzung des Gegenstandes bei Feststellen der Identität der Stimme mit einer solchen, deren Signale bereits in der Auswerteeinheit des Speichers abgelegt sind, möglich ist; bei Feststellen mangelnder Identität aber unmöglich ist. Dies kann beispielsweise durch einen Schalter in einem Stromkreis des Gegenstandes erfolgen, ohne dessen Geschlossensein der Gegenstand nicht in Benutzung oder in Betrieb genommen werden kann. Im Falle der Sicherung eines Kraftfahrzeugs kann es sich dabei um einen Schalter handeln, der beim Erkennen einer abgespeicherten Stimme die vorher gesperrte Bordelektrik oder -elektronik freigibt.

Das erfindungsgemäße Verfahren wird mit Hilfe einer Auswerte- und Steuereinheit durchgeführt, die z. B. aus einem ASIC bestehen kann. Diese Einheit steuert sowohl das Einlernen als auch das Vergleichen und Identifizieren der Stimme: in einer Aufnahmeeinheit, beispielsweise mit Mikrofon, wird die Stimme des Benutzers zuerst in elektrische Signale umgewandelt, evtl. bandgefiltert und dann digitalisiert. Das digitalisierte Signal wird an die Auswerteeinheit weitergegeben und dort verglichen. Bevorzugt wird dabei ein Vergleichsalgorithmus eingesetzt, der in der Auswerteeinheit abgespeichert sein kann. Dieser Vergleichsalgorithmus hat zum Ziel, die digitalisierten Daten einer sich identifizierenden Person mit den im System gespeicherten Daten anderer, beispielsweise zum Benutzen des Gegenstandes befugter Personen zu vergleichen und so zu entscheiden, ob es sich bei der sich identifizierenden Person um eine befugte Person handelt.

Der Vergleichsalgorithmus berechnet eine Short-Time-Fouriertransformation der eingehenden digitalisierten Daten. Der dabei entstehende neue Datenvektor (auch Merkmalsvektor genannt) wird mit Hilfe eines Klassifikationsverfahrens mit den abgespeicherten Daten verglichen. Als Klassifikatoren können statistische Klassifikationsverfahren, neuronale Netze oder synergetische Computer verwendet werden. Der Klassifikator entscheidet, ob der Merkmalsvektor mit dem Merk-

malsvektor einer gespeicherten Stimme so gut übereinstimmt, daß er diese Person als "im wesentlichen identisch" mit einer Person klassifiziert, deren Stimme gespeichert ist (Anspruch 7, Anspruch 8, Anspruch 9).

Wird das Verfahren zur Diebstahlsicherung eingesetzt, dann wird bei dieser — positiven — Klassifizierung die Diebstahlsicherung ausgeschaltet bzw. das "Schloß" mit der als "Schlüssel" erkannten Stimme aufgeschlossen. Im Falle eines Kfz wird dann beispielsweise die Bordelektronik freigegeben oder der Stromkreis zwischen Batterie und Zündung geschlossen.

Wird die Stimme als "fremd" klassifiziert, wird die Sperre nicht aufgehoben. Optional können optische oder akustische Warnsignale ausgegeben werden, beispielsweise mit Hilfe der Hupe oder Lichthupe (Anspruch 6).

Vorzugsweise wird die Spannungsversorgung des Sicherungssystems, über einen Akku gepuffert, von der Batterie des Kfz sichergestellt.

Ebenfalls optional kann ein Ausgang der Auswerteeinheit an vorhandene Bordlautsprecher (Radioanlage) abgeschlossen werden. Dies ermöglicht zusätzlich, daß ein Benutzer vom System durch einen akustischen Hinweis (z. B. den Satz "bitte identifizieren sie sich" oder einen Piepston) zum Versuch der Identifikation aufgefordert wird.

Es ist bevorzugt daß das "Einlernen" neuer befugter Benutzer nur bei ausgeschalteter Diebstahlsicherung erfolgen kann. Bevorzugt sind hierfür zwei Tasten im Fahrerraum vorgesehen, deren eine den "Einlernmodus" aktiviert. Ist dieser Modus aktiviert, kann der "neue Benutzer" mehrmals hintereinander charakteristische Laut- oder Tonfolgen von sich geben, beispielsweise seinen Namen oder ein frei wählbares Kennwort sprechen. Dieses Wort oder diese Lautfolge wird vom System aufgenommen und vom Klassifikator der Auswerteeinheit in einer geeigneten Form im Speicherspeicher angelegt. Anhand genau dieses Wortes wird der Benutzer nachher wieder identifiziert.

Bevorzugt erfolgt das Einschalten der Diebstahlsicherung von Kfz bei ausgeschalteter Zündung über eine weitere Taste im Fahrerraum. Ganz besonders bevorzugt ist es, daß die einmal eingeschaltete Sicherung nur durch das gesprochene Kennwort eines befugten Benutzers wieder deaktiviert wird.

Ein Beispiel soll das Verständnis der Erfindung vertiefen.

Fig. 1 zeigt das Beispiel einer Anordnung, mit der das erfindungsgemäße Verfahren durchgeführt werden kann, wenn ein Gegenstand wie ein Kfz gegen Diebstahl gesichert werden soll.

Im Bordcomputer befindet sich ein Chip 1 für die Datenauswertung (z. B. ein ASIC), der die Abspeicherung von Stimmen befugter Benutzer in einem Speicher 2 (z. B. einem EPROM oder EEPROM) steuert. Der Chip ist an die Stromversorgung der Batterie 6 angeschlossen. Die Versorgung kann über Akkus 3 gepuffert sein.

Ein Ausgang des Chips 1 ist mit einer Aufnahmeeinheit 4 mit einem Mikrofon verbunden. Ein weiterer Ausgang ist über einen Schalter 5 mit dem Stromkreis verbunden, der die Batterie 6 mit der Zündung oder Bordelektronik 7 verbindet.

Fakultativ sind Schalter oder Taster zum Scharfstellen der Diebstahlsicherung (T1) bzw. zum Starten des Einlernmodus für neue Benutzer (T2) vorgesehen. Ebenfalls fakultativ ist ein Ausgang des Chips mit einem Lautsprecher 8, beispielsweise dem Radio, verbunden,

durch welchen eine akustische Aufforderung zur Benutzeridentifikation aktiviert werden kann.

Die Klassifizierung des gesprochenen Wortes entscheidet, ob der über eine FFT gewonnene Merkmalsvektor der Sprache der Person, die das Kfz öffnen oder anlassen möchte, mit dem im Speicher 2 Vorgegebenen hinreichend übereinstimmt. Bejahendenfalls wird die Fahrertür geöffnet oder das Anlassen erlaubt.

Mit einer Direktanlassung kann ebenfalls gearbeitet werden, so daß Schlüssel entbehrlich sind. Das Abschalten des Motors dagegen kann durch einen weiteren optionalen Taster erfolgen, hierfür ist die Benutzung des Wortes (dessen Merkmalsvektor) nicht zwingend, vorteilhaft sogar entbehrlich.

Patentansprüche

1. Verfahren zum Identifizieren einer menschlichen Stimme, insbesondere zum Sichern von Gegenständen gegen Diebstahl, umfassend

- (a) das Aufnehmen einer Stimme mit Hilfe einer Aufnahmeeinheit (4),
- (b) das Umwandeln der aufgenommenen akustischen Signale in elektrische Signale (4),
- (c) das Digitalisieren und Abspeichern der Signale in einer Auswerteeinheit mit Speicher (1, 2),

wobei ggf. die Schritte (a) bis (c) mit weiteren Stimmen wiederholt werden, und

- (d) das Vergleichen digitalisierter Signale einer Vergleichs-Stimme mit einer oder mehreren solchen, die sich bereits im Speicher (2) der Auswerteeinheit (1) befinden, und Identifizieren der Stimme als im wesentlichen identisch mit der oder einer der Stimmen, die im Speicher (2) der Auswerteeinheit (1) gespeichert sind.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die Auswerteeinheit (1) über einen ersten Schalter (T2) so gesteuert wird, daß nur die Schritte (a) bis (c) ausgeführt werden können, und/oder über einen zweiten Schalter (T1) so gesteuert wird, daß nur Schritt (d) ausgeführt werden kann.

3. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß die Auswerteeinheit (1) so mit dem zu sichernden Gegenstand verbunden ist, daß die uneingeschränkte Benutzung des Gegenstandes bei Feststellen der Identität der Vergleichs-Stimme mit einer solchen, deren Signale bereits im Speicher (2) abgelegt sind, möglich ist, bei Feststellen mangelnder Identität unmöglich ist.

4. Verfahren nach Anspruch 3, dadurch gekennzeichnet, daß es sich bei dem zu sichernden Gegenstand um ein Kraftfahrzeug handelt, daß die Verbindung einen Schalter (5) umfaßt, der den zum Starten des Kraftfahrzeugs erforderlichen Stromkreis unterbrechen kann und daß die uneingeschränkte Benutzung das Starten und Wegfahren ist.

5. Verfahren nach Anspruch 2 und 4, dadurch gekennzeichnet, daß sich der erste und/oder der zweite Schalter (T1, T2) zur Steuerung der ausführbaren Schritte des Identifizierens im Innenraum des Kraftfahrzeugs befinden.

6. Verfahren nach einem der Ansprüche 3 bis 5, dadurch gekennzeichnet, daß beim Feststellen mangelnder Identität akustische oder optische Warnsignale abgegeben werden.

7. Verfahren nach einem der voranstehenden Ansprüche, dadurch gekennzeichnet, daß in Schritt (d) die Stimmen mit Hilfe eines Vergleichsalgorithmus verglichen werden, der eine Short-Time-Fouriertransformation (FFT) der eingehenden digitalisierten Daten berechnet, wobei der neu entstehende Datenvektor mit Hilfe eines Klassifikationsverfahrens mit Datenvektoren der abgespeicherten Signale verglichen wird. 5
8. Verfahren nach Anspruch 7, dadurch gekennzeichnet, daß für das Klassifikationsverfahren neuronale Netze oder synergetische Computer verwendet werden. 10
9. Verfahren nach einem der Ansprüche 3 bis 8, dadurch gekennzeichnet, daß der Klassifikator entscheidet, ob die verglichenen Merkmalsvektoren so gut übereinstimmen, daß im wesentlichen Identität festgestellt wird. 15
10. Vorrichtung zum Identifizieren einer menschlichen Stimme gemäß einem der erwähnten Verfahrensansprüche, umfassend 20
- (a) eine Aufnahmeeinheit mit Mikrofon (4),
 - (b) eine Auswerte- und Steuereinheit (1),
 - (c) einen Speicher (2) für digitalisierte Stimmsignale, sowie 25
 - (d) eine Verbindung zu dem zu sichernden Gegenstand, die die Inbetriebnahme des Gegenstandes steuern kann (5), und/oder
 - (e) einen oder mehrere Akkumulatoren (3, 6) für eine gepufferte Spannungsversorgung der Vorrichtung, und/oder 30
 - (f) einen oder mehrere Lautsprecher (8), und/oder
 - (g) Schalter (T1, T2) zum Steuern der Bedienungsfunktionen der Vorrichtung. 35
11. Verwendung mindestens eines gesprochenen Wortes als Schlüssel für das Öffnen einer Autotür oder als Schlüssel für das Freigeben des Motorstarts oder selbiges selbst. 40

Hierzu 1 Seite(n) Zeichnungen

45

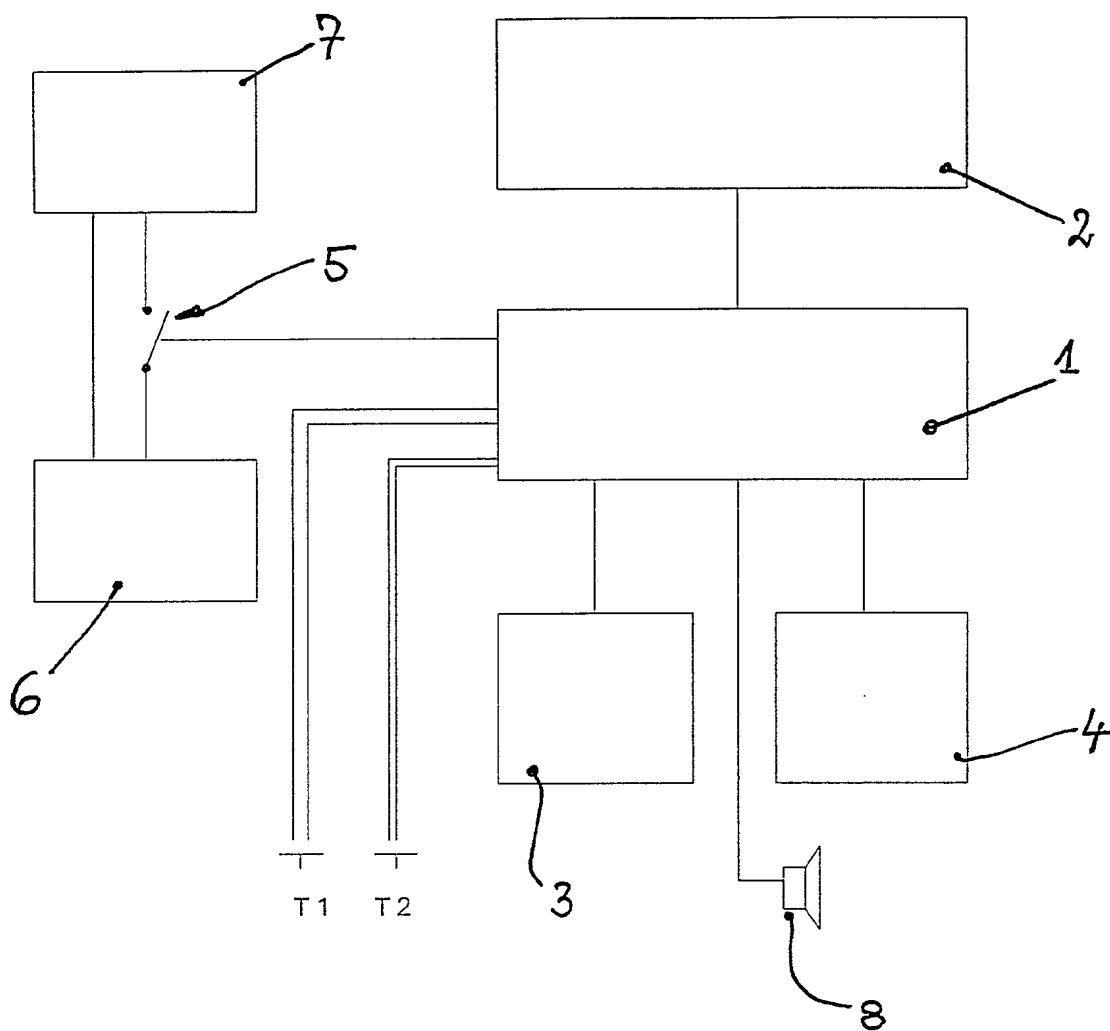
50

55

60

65

- Leerseite -



Figur 1